DOCUMENT RESUME

ED 161 005 CS 004 439

| | |
|---|---|
| TITLE | Testing and Evaluation in Reading: Abstracts of Doctoral Dissertations Published in "Dissertation Abstracts International," May through September 1978 (Vol. 38 No. 11 through Vol. 39 No. 3). |
| INSTITUTION | ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill. |
| PUB DATE | 78 |
| NOTE | 11p. |
| EDRS PRICE | MF-$0.83 HC-$1.67 Plus Postage. |
| DESCRIPTORS | Adults; Annotated Bibliographies; Content Reading; Criterion Referenced Tests; Critical Reading; *Doctoral Theses; Elementary Secondary Education; *Learning Disabilities; Miscue Analysis; *Predictive Ability (Testing); Readability; Reading Achievement; Reading Diagnosis; Reading Readiness; *Reading Research; Reading Skills; *Reading Tests; Spelling; Test Validity; Word Recognition |

ABSTRACT

This collection of abstracts is part of a continuing series providing information on recent doctoral dissertations. The 13 titles deal with the following topics: the evaluation of the elementary reading specialist; the validity of the Visual-Aural Digit Span Test as a predictor of word recognition and spelling in learning disabled children; the validation of three tests devised to measure critical reading ability; a criterion-referenced test of fundamental reading competencies; the influence of reading behavior procedures on reading test performance; kindergarten-first grade screening measures as predictors of third grade reading achievement; differences in performance on the traditional and object form of a reading readiness test; the origin and development of reading diagnosis in the United States from 1896 to 1946; a comparison of two methods of helping content-area teachers select and use student reading materials; problem solving behavior of teachers diagnosing a child's reading performance; an unobtrusive screening device to approximate the reading levels of adults; the relationship between reading miscue inventory scores and readability level; and criterion-referenced word recognition measures as predictors of reading achievement. (HAI)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED161005

# Testing and Evaluation in Reading:

Abstracts of Doctoral Dissertations Published in Dissertation Abstracts International, May through September 1978 (Vol. 38 No. 11 through Vol. 39 No. 3)

Compiled by the Staff of

The ERIC Clearinghouse on Reading and Communication Skills

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

University Microfilms
International

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

CS004439

The dissertation titles contained here are published
with permission of the University Microfilms International,
publishers of Dissertation Abstracts International
(copyright © 1978 by University Microfilms International),
and may not be reproduced without their prior permission.

This bibliography has been compiled as part of a continuing series designed to make information on relevant dissertations available to users of the ERIC system. Monthly issues of Dissertation Abstracts International are reviewed in order to compile abstracts of dissertations on related topics, which thus become accessible in searches of the ERIC data base. Ordering information for the dissertations themselves is included at the end of the bibliography.

Abstracts of the following dissertations are included in this collection:

Button, Norma J.

A PARADIGM TO GUIDE THE EVALUATION OF THE ELEMENTARY READING SPECIALIST

Coff, Laurence A.

THE VALIDITY OF THE VISUL-AURAL DIGIT SPAN TEST AS A PREDICTOR OF WORD RECOGNITION AND SPELLING FOR LEARNING-DISABLED CHILDREN

Dole, Janice Arnold

A VALIDATION OF THE CONSTRUCT OF CRITICAL READING AND OF THREE TESTS DESIGNED TO MEASURE CRITICAL READING ABILITY

Ganopole, Selina Judy

THE CONSTRUCTION AND VALIDATION OF A CRITERION-REFERENCED TEST OF FUNDAMENTAL READING COMPETENCIES

Homer, Cynthia Louise

THE INFLUENCE ON READING BEHAVIOR OF TWO PURPOSE-SETTING PROCEDURES FOR ADMINISTERING AN INFORMAL READING INVENTORY

Huffman, Robert George

KINDERGARTEN-FIRST GRADE SCREENING MEASURES AS PREDICTORS OF THIRD GRADE READING ACHIEVEMENT

Hutsell, Melba Lilly

DIFFERENCES IN PERFORMANCE ON THE TRADITIONAL AND OBJECT FORM OF A READING READINESS TEST

Pelosi, Peter Leon

THE ORIGIN AND DEVELOPMENT OF READING DIAGNOSIS IN THE UNITED STATES: 1896-1946

Rabin, Annette Tucker

A COMPARISON OF TWO METHODS OF IN-SERVICE PROGRAMMING FOR HELPING CONTENT-AREA TEACHERS SELECT AND USE STUDENT READING MATERIALS

Stephens, Elaine C. Amon

A STUDY OF THE PROBLEM-SOLVING BEHAVIOR OF TEACHERS AS THEY DIAGNOSE A CHILD'S READING PERFORMANCE AND THE EFFECTS OF EXPERIENCE AND TRAINING ON THAT BEHAVIOR

West, Ellen Margaret

THE USE OF AN UNOBTRUSIVE SCREENING DEVICE TO APPROXIMATE READING LEVELS OF ADULTS

Wolf, Anne Elizabeth

GENERALIZABILITY OF READING MISCUE INVENTORY SCORES GENERATED BY AVERAGE READERS IN GRADE THREE ON BASAL-READER FICTIONAL SELECTIONS OF COMPARABLE LENGTH AND READABILITY LEVEL

Wright, Dennis Albert

CRITERION-REFERENCED WORD RECOGNITION MEASURES AS PREDICTORS OF READING ACHIEVEMENT

A PARADIGM TO GUIDE THE EVALUATION OF THE ELEMENTARY READING SPECIALIST Order No. 7814541

BUTTON, Norma J., Ph.D. Saint Louis University, 1978. 112pp.

The purpose of the study was to create a paradigm which could serve as a model for evaluation of the elementary reading specialists by attempting to define the competencies of reading specialists using a systematic procedure, and determine the fit of identified qualities of reading specialists to evaluation instruments.

A group of Greater St. Louis area reading personnel identified quality components regarding the activities of reading specialists. A second group of reading personnel in the Parkway School District assisted in the development of values-based criteria for the appraisal of the elementary reading specialist.

These suggestions were combined with information gained from a survey of the literature to create a paradigm to lead the development of instruments to evaluate elementary reading specialists. The paradigm was not intended as a quantitative measure of appraisal, but rather an organization for consideration of the scope and comprehensiveness of the role of the elementary reading specialist for a model from which evaluation instruments could be developed.

A structured telephone interview guide was created and utilized in conversation with the local St. Louis area reading personnel in order to gain information concerning quality components, necessary competencies, and existing evaluation instruments. These conversations were taped, with permission, and studied for further insights. In addition, written information was obtained through a negotiation process with Parkway reading personnel regarding concrete quality activities performed by reading specialists. These insights were gained through individual introspection and small group compilation.

When the sets of data were carefully scrutinized, a very close accord was found between the information obtained from the local St. Louis area reading personnel and that of Parkway reading personnel, giving much credence to the hypotheses that reading personnel are able to identify quality components regarding their activities and can indeed assist the development of values-based criteria for the appraisal of the elementary reading specialist, and that a meaningful paradigm leading to the creation of instruments for evaluating reading specialists using a perception-based model is possible. The processes involving focusing the perceptions, describing the perceptions, and judging the perceptions occurred confluently with each other and were taken as a total to form the theoretical basis for the evaluation paradigm.

Upon completion, the paradigm was presented to a judgment team, comprised of six specialized personnel in the field of education. The judgment team was responsible for review and validation of the paradigm.

THE VALIDITY OF THE VISUAL-AURAL DIGIT SPAN TEST AS A PREDICTOR OF WORD RECOGNITION AND SPELLING FOR LEARNING-DISABLED CHILDREN Order No. 7732426

COFF, Laurence A., Ph.D. California School of Professional Psychology, San Diego, 1977. 220pp.

The purpose of this study was to determine the validity of the Visual-Aural Digit Span (VADS) Test as a predictor of word recognition and spelling for dyslexic children. It was hypothesized that the VADS Test, which purportedly measures immediate memory, sequencing, and visual-aural integration, would be significantly correlated with word recognition and spelling, as measured by the Wide Range Achievement Test (WRAT). It was also hypothesized that improvement in the VADS Test would be significantly correlated with improvement in spelling and word recognition. Improvement measures were obtained by comparing pretest and posttest scores on test batteries which were administered approximately eight months apart. Forty-five subjects from year-round learning disability classrooms participated in the pretest portion of this study. Thirty-four subjects took both the pretest and posttest battery.

The first two hypotheses, that the VADS Test would be significantly correlated with the reading and spelling subtests of the Wide Range Achievement Test, were supported by the data. However, the hypotheses that VADS Test improvement would be significantly correlated with improvement in spelling and word recognition were not supported by the results of this study. In addition, t-tests comparing the pretest and posttest criterion variables found that a significant improvement occurred in spelling and word recognition. However, there was no significant improvement in the VADS Test.

Several limitations on the construct validity of the VADS Test were discussed as possible explanations of these results. In addition, the implications of these results for information processing theory and the maturational lag theory were elaborated. However, the most important finding of this study is that reading and spelling improvement can occur without a corresponding improvement in immediate memory, sequencing, and visual-aural integration, as measured by the VADS Test. Further research is recommended on the use of the VADS Test as a screening instrument for the early detection of dyslexia.

A VALIDATION OF THE CONSTRUCT OF CRITICAL READING AND OF THREE TESTS DESIGNED TO MEASURE CRITICAL READING ABILITY Order No. 7808892

DOLE, Janice Arnold, Ph.D. University of Colorado at Boulder, 1977. 193pp. Director: Professor Donald E. Carline

Although the specific skills and abilities involved in the reading comprehension process have not been conclusively identified, general agreement exists that they involve, in part, the ability to derive information explicitly stated in a passage and the ability to interpret information at "the higher levels of mental functioning" (Sochor, 1958). The former ability has been labeled literal reading ability; the latter has been designated for this study as critical reading ability.

In contrast to the construct of literal reading, the construct of critical reading has been defined in numerous ways and has been identified through a broad range of skills and abilities. In an effort to clarify this one aspect of reading comprehension, a construct validation of critical reading was undertaken. A multitrait-multimethod approach (Campbell and Fiske, 1959) was used to determine if critical reading could be differentiated from the theoretically related construct of literal reading.

Measures of the constructs in different content areas were

employed to evaluate validity and to examine possible content area influences on the constructs. Two critical reading tests, one in science and one in social studies, were selected from the literature. An additional, parallel critical reading test in English literature was constructed by the author. The CTBS reading comprehension test, previously administered to the subjects used in this study, was broken down by content area and reading level and employed as an additional method of measurement. A teacher rating instrument developed by the author comprised the final method of evaluation.

The subjects for the study were 136 volunteer sixth graders drawn from four schools in a predominantly white, middle to upper-middle class community.

Results of the study indicated modest convergent validity for both literal and critical reading across the content areas. English literal reading and social studies critical reading averaged the highest (.50 and .52 respectively) validity values. Average convergent validity values were .41 for both literal and critical reading. While these correlations are statistically significant, they represent only moderate evidence that the various instruments are measuring the same constructs.

No discriminant validity was found for the constructs in different content areas. Some evidence was found to suggest that English and science literal reading could be differentiated as unique constructs. This trend manifested itself when the content areas were combined to form a new matrix. In the combined content area matrix literal reading had a higher convergent validity value (.61) and showed some evidence of discriminant validity. Critical reading, however, demonstrated no uniqueness as a construct in either matrix.

Results of the study indicated that the critical reading instruments in different content areas were better measures of reading ability in general than of critical reading ability in particular. More valid instruments need to be developed to measure critical reading. In addition, alternative definitions of critical reading need to be examined and would perhaps be more useful in a description of reading comprehension than would the operationalization of critical reading used in this study.

THE CONSTRUCTION AND VALIDATION OF A CRITERION-REFERENCED TEST OF FUNDAMENTAL READING COMPETENCIES Order No. 7811349

GANOPOLE, Selina Judy, Ed.D. University of California, Los Angeles, 1978. 217pp. Chairman: Professor W. James Popham

As a result of the minimum competency testing movement in California, local school districts are required to adopt measurable standards of proficiency in the basic skill areas of reading comprehension, writing, and computation. While large numbers of reading achievement tests are available, the majority of these may be inappropriate to a search for minimal competency assurance for all high school students. Further, recent use of the term "minimal competencies" as it relates to reading, connotes the ability to perform functional reading tasks, that is, reading tasks which are essential for day-to-day living and effective participation in society. The concept that functional reading directly relates to reading experiences encountered in day-to-day living is not typically reflected in the content or format of traditional tests of reading achievement.

As a response to districts' needs to assess student proficiency with respect to functional reading skills and the unavailability of appropriate assessment instruments, this study focused on the development and validation of such a test. The Fundamental Reading Competencies Test (FRCT) was designed to measure the ability of high school students to meet the reading requirements encountered in normal everyday living. Mastery of the competencies selected for inclusion in this test was considered essential for successfully meeting the reading requirements of daily living in contemporary society.

The study progressed through three principal stages: the selection of reading competencies to be assessed, the construction of test items to measure those competencies, and the establishment of the validity and reliability of the instrument.

The initial stage, selection of competencies, involved the identification of specific, assessable reading competencies. An extensive pool of tentative competencies was gathered by drawing upon such sources as the National Assessment of Educational Progress, the Adult Performance Level Study, the Educational Testing Service studies of adult reading behaviors, extant collections of reading objectives, as well as competency requirements in other states and current research in the field of functional reading. A major focus during this stage centered around establishing a viable strategy for determining which competencies, from the universe of plausible reading competencies, should be selected for assessment. A related concern was the issue regarding the number of competencies which should be assessed by the test.

The second stage of the study centered around the construction of test items to assess the competencies selected. However, preliminary to the writing of the test items, it was necessary to define the limits of the behaviors that the test items would measure. This was accomplished by means of developing a set of test specifications for each of the competencies being assessed. Test specifications represent the set of rules needed to generate the items for a criterion-referenced test and, as such, are fundamental to any further development of the test.

Establishing the reliability and validity of the instrument constituted the major focus of the third stage of the study. Several strategies were employed to establish the test's validity, with special emphasis given to verifying the test's descriptive validity and its domain-selection validity. Three procedures were used to assess the reliability of the instrument: a test-retest reliability procedure, an equivalent forms reliability procedure, and a combination of these two procedures in order to obtain a stability and equivalence reliability estimate. In assessing the reliability of the test, reliability was considered and calculated from two rather distinct approaches: (1) the traditional, correlational approach typically used in computing estimates of norm-referenced measures and (2) a decision-consistency approach, suitable for use with criterion-referenced measures. Reliability estimates were calculated for each of the subtests as well as the test as a whole.

## THE INFLUENCE ON READING BEHAVIOR OF TWO PURPOSE-SETTING PROCEDURES FOR ADMINISTERING AN INFORMAL READING INVENTORY Order No. 7811178

HOMER, Cynthia Louise, Ed.D. Northern Illinois University, 1977. 288pp.

The intent of this investigation was to determine if individually formulated pupil purposes for reading, elicited by the examiner's higher level questioning strategy, and a priori purposes for reading, stated in the examiner's manual, would make a difference in reading performance on an informal reading inventory (IRI).

It was hypothesized that there would be no differences in: (1) independent, instructional, frustration reading levels and the range of levels between independent and frustration levels; (2) literal comprehension, word recognition, reading rate, and interpretive comprehension Standard Reading Inventory (SRI) subtest scores at the instructional reading level; and (3) quantitative and qualitative oral reading miscue patterns at the instructional reading level when an individually formulated purpose-setting procedure and an a priori purpose-setting procedure were used in the administration of the SRI.

In two separate individual testing sessions, 38 fourth grade students of average reading ability read two forms of the SRI. The purpose-setting procedures were alternated for the two SRI forms. In the experimental procedure, students were directed to: (a) formulate their own questions, derived from their predictions about the passage outcomes and the examiner's higher level questions; and (b) read the passage to answer their questions. In the control procedure, students were directed to read the passage and retell it at the conclusion of the reading.

The examiner scored and evaluated students' reading performances. To validate the results, two judges' evaluations of ten randomly selected protocols were compared to the examiner's evaluations. Mean percentages of inter-judge agreement were found to be high ($\bar{x}$ = .95 and .92).

The data were analyzed by two-tailed correlated t tests to determine significant differences between the two purpose-setting procedures. The level of significance for rejecting the null hypotheses was set at 0.05. Conclusions stated below were drawn from the analyses of the data.

When reading was directed by individually formulated pupil purposes, the reading achievement levels (independent, instructional, frustration) (a) were significantly higher, and (b) spanned a significantly higher and wider range of grade levels on an IRI than when reading was directed by a priori purposes. Average fourth grade readers who read for individual purposes also received a greater number of aided pronunciation cues at the instructional reading level.

When reading was directed by a priori purposes, average fourth grade readers (a) achieved significantly higher comprehension scores at the instructional reading level on SRI literal comprehension subtests of total retelling, total recall, silent retelling and recall, and oral retelling; and (b) achieved a significantly higher oral reading rate at the instructional reading level than when reading was directed by individually formulated pupil purposes.

When reading was directed by either purpose-setting procedure at the instructional reading level, average fourth grade readers (a) achieved equally well on SRI subtests of oral recall comprehension, word recognition, silent reading rate, and interpretive comprehension; and (b) generated about the same numbers of substitution, omission, addition, mispronunciation, word order, punctuation, and self-correction miscues which represented similar patterns of their use of graphic cues and contextual utilization cues. At the instructional reading level, silent reading rate was significantly higher than oral reading rate regardless of the two purpose-setting procedures utilized.

This study made a contribution to the information regarding the influence of purpose-setting behavior on reading performance in informal diagnostic assessment. The conclusions of the study imply that an increased emphasis on higher level questioning strategies in instructional settings should promote the development of critical reading and thinking processes and general reading achievement.

## KINDERGARTEN-FIRST GRADE SCREENING MEASURES AS PREDICTORS OF THIRD GRADE READING ACHIEVEMENT Order No. 7815322

HUFFMAN, Robert George, Psy.D., Rutgers University The State University of New Jersey, G.S.A.P.P., 1977. 143pp. Chairperson: Virginia D. C. Bennett

The purpose of this research was to study the effectiveness of an ongoing Early Identification Program including the following aspects: (a) its overall effectiveness in identifying those children likely to experience difficulty in learning to read, (b) the contribution of each screening instrument to the usefulness of the whole, (c) the manner in which the screening profile was used to determine those children in need of intervention, and (d) the criteria used by kindergarten and first grade teachers in making judgments about their pupils' probable future success in reading.

The sample consisted of 104 first grade boys and girls drawn from four elementary schools in an affluent Long Island community. Children who were not present for all screening and criterion measures, who exhibited significant sensory deficits, or who had repeated first or second grade were not included in the study.

Screening measures administered by district personnel during the sample's kindergarten and first grade years included vision and hearing screening, teacher judgments of children, the Metropolitan Readiness Tests, the Wepman Auditory Discrimination Test, and the Bender-Gestalt Test. The criterion measure, administered the Spring of third grade, was the Stanford Achievement Test in reading. The effects of I.Q. were controlled for through the use of Otis-Lennon group I.Q. scores.

The data were analyzed through the use of various descriptive statistics, as well as the SPSS computer programs for Pearson Correlation, Partial Correlation, and Regression.

Results of the study demonstrated that global measures of reading readiness and general ability, such as the Metropolitan, Otis-Lennon, and Teacher Judgment are much more effective predictors of later reading achievement than are tests of specific abilities such as the Wepman and Bender. However, the multiple correlation of these tests left nearly half of the variance unexplained, indicating that other unspecified factors are also important to reading success.

The findings further indicated that reading achievement is more readily predicted for girls than for boys. Additionally, it appears to be more difficult to predict reading achievement for children who experience early reading difficulties than for those who do not. I.Q. scores are also less effective predictors of achievement for children who demonstrate reading disabilities.

Other findings underlined the teacher's important role, both in identifying children with potential reading disabilities and in determining which children will receive help.

The implications of this study are that substantial cost and professional time could be saved by using global, group-administered screening measures, at least initially. Individual instruments might be administered to gain further information regarding a smaller number of identified youngsters.

Additionally, the findings suggest that it might be possible to increase the accuracy of early identification through greater emphasis on structured teacher observations and judgment. Related research findings suggest that inclusion of behavior ratings by teachers could further increase the accuracy of predictions.

## DIFFERENCES IN PERFORMANCE ON THE TRADITIONAL AND OBJECT FORM OF A READING READINESS TEST

Order No. 7815365

HUTSELL, Melba Lilly, Ph.D. The University of Oklahoma, 1977. 91pp. Major Professor: Robert L. Curry

This study assessed the level of individual pupil readiness for formal instruction in reading using object forms and pencil-paper tests to determine if some children are limited in answering questions on tests by the format of presentation rather than by limited knowledge concerning the correct answer.

First grade children were classified into groups according to their socioeconomic status and gender. The Lee-Clark Reading Readiness Test (1962 Revision) and an object form of the identical test material were administered individually. Samples of thirty-five males and thirty-five females from each of the upper, middle and lower socioeconomic levels were randomly drawn. Test items were alternately presented to the groups using the individually administered pencil-paper materials, object form test materials and six Piagetian Conservations Tools.

The results of testing the hypotheses regarding the score of both sex groups indicated that both boys and girls at each socioeconomic level performed slightly better on the pencil-paper version of the Lee-Clark Reading Readiness Test when it is administered individually than on the object form of the test materials. Each socioeconomic group performed slightly better on the pencil-paper version of the test when individually administered than on the object form of the test materials. Though slight gains were shown from the object form presentation of the materials to the individual pencil-paper presentation, it was determined that no significant differences exist between the performance of first grade students from different socioeconomic levels on the individual administration of the Lee-Clark Reading Readiness Test and an object form of the identical test materials. The number of conservers was so small that only the descriptive data regarding the number of conservers and non-conservers were reported.

## THE ORIGIN AND DEVELOPMENT OF READING DIAGNOSIS IN THE UNITED STATES: 1896–1946 Order No. 7813986

PELOSI, Peter Leon, Ed.D. State University of New York at Buffalo, 1977. 435pp.

This historical account of the inception and refinement of reading diagnosis traces its growth from the 19th Century through 1946. Three basic questions served as the basis for the research: 1. What were the influences and effects of the measurement movement upon the origin and development of reading diagnosis? 2. How did the implied and stated models or outlines of diagnostic procedures influence and relate to the development of this field in education? 3. What were the influences of authoritative reading opinions and changing ideologies in the field of reading diagnosis?

The growth and development of educational measurement from 1860 through 1915 was discussed as it eventually influenced the initial stages of reading diagnosis in the United States. Stages in the development of reading diagnosis were divided into the time periods of 1896–1925 which characterized the initial development of reading diagnosis. The years 1926–1932, the second stage, noted the development of the case study or clinical approach to diagnosing reading problems. In the third stage, 1933–1938, there was discussed the search for multiple causes and the beginnings of standardization in the diagnostic process. From 1939–1946 the final stage of reading diagnosis was characterized by the development of the informal reading inventory and development of a precise interdisciplinary approach to this specialized area in education. Conclusions covering the span of time the history covered are presented in the seventh and final chapter.

## A COMPARISON OF TWO METHODS OF IN SERVICE PROGRAMMING FOR HELPING CONTENT-AREA TEACHERS SELECT AND USE STUDENT READING MATERIALS

Order No. 7815953

RABIN, Annette Tucker, Ed.D. University of Pennsylvania, 1978. 240pp. Chairman: Dr. Allan Glatthorn

The purpose of this study was to compare the effectiveness of two types of in-service programing in changing the attitude and behavior of content area teachers toward the use of reading techniques in their classrooms while at the same time providing useful assistance to those teachers involved.

An in-service workshop on the selection and adequate reinforcement of reading materials used in the content area classroom was held for the entire faculty of a junior high school approximately 20 miles west of Philadelphia. The workshop, conducted by the school's reading department, was in two sections: a large group session during which teachers were taught to use the Fry Readability Graph and interpret standardized reading scores so that they might compare the reading levels of the students in their classes with the readability levels of classroom written materials and three simultaneous follow-up sessions during which smaller groups of teachers were instructed in ways of revising and/or reinforcing existent written materials. Each participant was given a set of materials prepared by the investigator which included a master packet on readability and four related "activity" packets.

Twenty-nine teachers of English, social studies, mathematics, and science were then divided into experimental and control groups for an action research study. Treatment for the controls consisted of the workshop and related packets only. Treatment for the experimental group consisted of the above plus individualized assistance with actual classroom situations and weekly observations over a three month period.

An attitude inventory, devised by Wayne Otto, was used as a pre- and post-test to measure change in the attitudes of the content area teachers toward the use of reading techniques in their classes. Behavior change was measured by three instruments developed by the investigator: the Post Observation Interview Form, completed after each classroom observation, the Faculty Survey, designed to determine to what extent either group had used the information from the workshop and packets, and a Post Summer Survey, administered the following fall to determine the effect of a longer period of time and more leisure upon the use of the information.

At the end of the study, both groups were more positive in their attitudes toward the integration of reading techniques into their particular content areas. Those in the experimental group, however, unlike the controls, perceived themselves as being capable of using reading techniques in their classes without formal instruction and also saw more possibility for self-satisfaction and excitement in teaching the necessary skills to their students than did controls.

During the observation period neither group of teachers noticeably changed its methods of choosing and presenting written materials. Both groups gave the same major reasons for not following through with the techniques introduced in the readability workshop: inadequate preparation time and a necessity for retaining technical vocabulary which precluded reducing the reading levels of certain materials.

The intervention of a summer did not make much difference in behavior. Teachers in the experimental group had made more use of the workshop packets, mostly for computing the readability levels of new materials. New learning packets produced by teachers independently or under district supervision continued to be one track only.

Though members of the experimental group had not changed their behaviors greatly on their own, they were most interested in continued assistance from the reading department in evaluating written materials, developing curricula on lower readability levels, identifying students reading appreciably above and below grade level, and in general, following through with strategies and techniques introduced as part of the in-service program. Controls, also, expressed interest in continued assistance but to a lesser degree.

A STUDY OF THE PROBLEM-SOLVING BEHAVIOR OF TEACHERS AS THEY DIAGNOSE A CHILD'S READING PERFORMANCE AND THE EFFECTS OF EXPERIENCE AND TRAINING ON THAT BEHAVIOR Order No. 7815171

STEPHENS, Elaine C. Amon, Ph.D. Michigan State University, 1978. 130pp.

It appears to be an established principle of the reading profession that classroom teachers should be diagnostic teachers of reading. Not so well established, however, are principles to help answer these two major concerns: How are reading problems diagnosed? and How do teachers learn to diagnose reading problems? What teachers actually do when they diagnose is largely speculative. Specifically, how a teacher uses her problem-solving skills to diagnose a child's reading performance remains unanswered. Therefore, this study attempted to (1) identify and describe the problem-solving behavior of teachers as they diagnosed a child's reading problem, and (2) determine the effects of classroom teaching experience and graduate level instruction in reading diagnosis and correction on that behavior.

A review of the literature indicated the lack of a sufficient model of problem-solving behavior for reading diagnosis based on empirical evidence. Due to the insufficiency of present models, this study, as well as the larger research project of which it is a part, used a model developed in another discipline which attempts to describe diagnostic problem-solving behavior and which may have application to reading diagnosis. Designated the Inquiry Theory of Clinical Problem Solving, it describes diagnostic problem-solving as a complex form of reasoning which is probabilistically determined by the interplay of the problem-solver's previously acquired cognitive capabilities for the problem and specific properties of the problem itself. It suggests that the problem-solver's performance is a function of his store of knowledge about and search strategies for a particular problem. If these can be aided or improved, then diagnostic performance should improve. Previous research in other disciplines seems to indicate that experience and training in diagnostic problem solving improve knowledge and strategy thus resulting in improved performance.

This study attempted to apply the Inquiry Theory of Clinical Problem Solving to teachers and their diagnosis of a child's reading problem by hypothesizing that graduate level instruction in reading diagnosis and correction and/or classroom teaching experience would improve teachers' knowledge about reading problems and their search strategies for collecting and processing information, thus resulting in improved diagnostic performance. To this end, thirty pre-service and classroom teachers were assigned to three groups each containing ten subjects. The criteria for selection and assignment to groups were classroom teaching experience and graduate level instruction in the diagnosis and correction of reading difficulties. Under the constraints of procedures which were designed to elicit problem-solving behavior, each teacher interacted with a set of materials which simulated the behavior of a child with a problem in reading.

Written protocols and tape recordings of the teacher's behavior during the interaction as well as statements of the final diagnosis and remediation plan were analyzed to obtain data for designated measures of problem-solving behavior. Both product and process measures of diagnostic performance were used. Product measures were used to measure the final stated diagnosis of the reading problem. Process measures were used to describe the manner in which the problem was diagnosed. One-way multivariate analysis of variance and Pearson Product-Moment Correlation were the statistical procedures employed.

The major finding of this study was that there was no significant mean score difference on the principal product measure of diagnostic performance between teachers with teaching experience and graduate instruction in reading diagnosis and correction, teachers with teaching experience and no graduate instruction in diagnosis, and pre-service teachers without teaching experience or graduate instruction. Given the limitations of the present study, it appears that neither graduate level instruction in reading diagnosis and correction and/or classroom teaching experience seemed to significantly affect the knowledge and strategies employed by teachers in diagnosing a child's reading problems under simulated conditions and, therefore, did not result in improved diagnostic performance.

THE USE OF AN UNOBTRUSIVE SCREENING DEVICE TO APPROXIMATE READING LEVELS OF ADULTS
Order No. 7810996

WEST, Ellen Margaret, Ed.D. The University of Florida, 1977. 79pp. Chairman: Dr. Ruthellen Crews

The need for a minimal threat, reading evaluation instrument has been recognized by adult educators for some time. There have been several attempts to devise such an instrument; however, all of these attempts have utilized formal testing situations and have often required the adult student to read orally.

The purpose of this study was to develop an instrument which could serve as an unobtrusive screening device to approximate adult reading levels prior to class assignment, and, to test the validity of this instrument as well as the reliability of its scores.

The instrument, known as the West Informal Reading Evaluation (WIRE) was designed to fit the format of a student survey and was presented in three separate passages. These passages were written at the third, sixth, and ninth grade level with the use of the Maginnis (1969) adapted formula for use with the Fry Readability Graph (1968). Each passage dealt with a topic relevant to adults and specific to the school in which the study took place.

The study sample consisted of 154 adult education students, who completed form-1 of the WIRE. One-hundred twenty-three of these same students also completed the Reading for Understanding Placement Test (RFU) (Thurstone, 1963), which was used as the validity criterion in the study. In order to test reliability of the WIRE scores, 100 students of the original 154 were also asked to complete a parallel form of the WIRE, the WIRE-2.

The data were analyzed using the Pearson r Formula. Results showed a significant correlation between the WIRE-1 and RFU with a validity coefficient of .866. Validity for the WIRE-2 was then determined using the same formula and resulted in a coefficient of .818, which was also significant at the .01 level.

Reliability was determined using the Pearson r Formula as applied to the scores of the WIRE-1 and the WIRE-2. The result was a reliability coefficient of .847, which was significant at the .01 level.

Subjective data collected by the WIRE were also noted and were of considerable value. Summaries of passage responses were included in the study.

Findings of the study showed that the WIRE was a valid and reliable instrument that could be used as an unobtrusive method of informal reading evaluation.

## GENERALIZABILITY OF READING MISCUE INVENTORY SCORES GENERATED BY AVERAGE READERS IN GRADE THREE ON BASAL-READER FICTIONAL SELECTIONS OF COMPARABLE LENGTH AND READABILITY LEVEL

Order No. 7811755

WOLF, Anne Elizabeth, Ph.D. The University of Wisconsin – Madison, 1978. 263pp. Supervisor: Professor Wayne Otto

The primary purpose of this study was to investigate the reliability of Reading Miscue Inventory (RMI) scores when subjects read different reading selections. Reliability was investigated within the framework of Cronbach's Generalizability Theory. This liberalization of traditional reliability theory explicitly accounts for multiple sources of variation in observed test scores. The one facet, or factor, design of the study allowed for estimation of variance components for subjects, reading selections, and unexplained variation. Rater, administrator, and transcriber were held constant, paralleling many real-world uses of the RMI.

Twenty average readers in grade three read three fictional basal-reader selections of comparable length written at the grade four readability level. Univariate analyses of variance were calculated for the nine RMI categories and for four scoring options; these paired scores for either the first 25 errors on a selection or all the errors made with one of two dependent variables: either highly similar/acceptable errors only or highly and partially similar/acceptable errors combined. Results were reported for all RMI categories except dialect, for which too few errors were generated.

The secondary purpose of the study was to investigate three methodological questions: Would the number of errors scored affect the outcome of the study? Would the choice of dependent variable affect the outcome of the study? Would the three selections written at the same readability level be of comparable difficulty according to three other measures?

The following conclusions were drawn:

1. Generalizability coefficients in virtually all RMI categories under all four scoring options were too low to justify using scores for decision-making in research or diagnosis.
2. More variance in observed RMI scores was unexplained than explained. Facets not included in the study, such as rater and subject-selection interaction, are likely contributors to this unexplained variation.
3. Subjects varied considerably in ability to use graphophonic, syntactic, and semantic constraints when making oral reading errors.
4. There was considerable constant error associated with reading different selections.
5. The generalizability coefficients for RMI scores based on only the first 25 errors were much lower than those based on all the errors made.
6. In RMI categories with three scoring alternatives, generalizability coefficients differed when different univariate dependent variables were created, but there was no obvious pattern to the affect.
7. Matching selections on readability level did not ensure that they were of comparable difficulty on other measures.
8. The concept of passage difficulty is viable: selections could be rank-ordered for difficulty.

## CRITERION-REFERENCED WORD RECOGNITION MEASURES AS PREDICTORS OF READING ACHIEVEMENT

Order No. 7810567

WRIGHT, Dennis Albert, Ph.D. Arizona State University, 1978. 123pp.

The purpose of this study was to determine if criterion-referenced word recognition measures were adequate predictors of reading achievement. Literature was reviewed to identify a valid measure of reading achievement, to describe the characteristics of useful criterion-referenced tests, and to define the related elements of the word recognition process.

The study involved 119 third graders in two reading measures. The first was a grade level cloze passage constructed according to accepted procedure. The second measure was a criterion-referenced word recognition test consisting of three subtests: Auditory Phonics, Visual Phonics, and Structural Analysis.

The statistical relationship between the two reading measures was examined through regression analysis. Separate word recognition sub-test measures were found inadequate as predictors of the cloze reading achievement measures. Various combinations of the word recognition measures were also found inadequate as predictors of the cloze reading achievement measures.

Limited conclusions were drawn concerning the relationship between the cloze reading achievement measures and the criterion-referenced word recognition measures. The criterion-referenced word recognition measures were not adequate predictors of overall reading achievement as determined by the cloze procedure. Statistical regression techniques have questionable application with criterion-referenced reading measures.

Research recommendations were made in several areas. Format and content of criterion-referenced reading measures need empirical validation. Alternatives to regression techniques in the statistical analysis of criterion-referenced reading measures need evaluation. The effect of repeated applications of the cloze procedure and variations of introductory tasks on cloze scores require further investigation.

Copies of the dissertations may be obtained by addressing your request to:

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106

or by telephoning (toll-free) 1-800-521-3042